Update on Sparse CNNs for Particle ID in ProtoDUNE

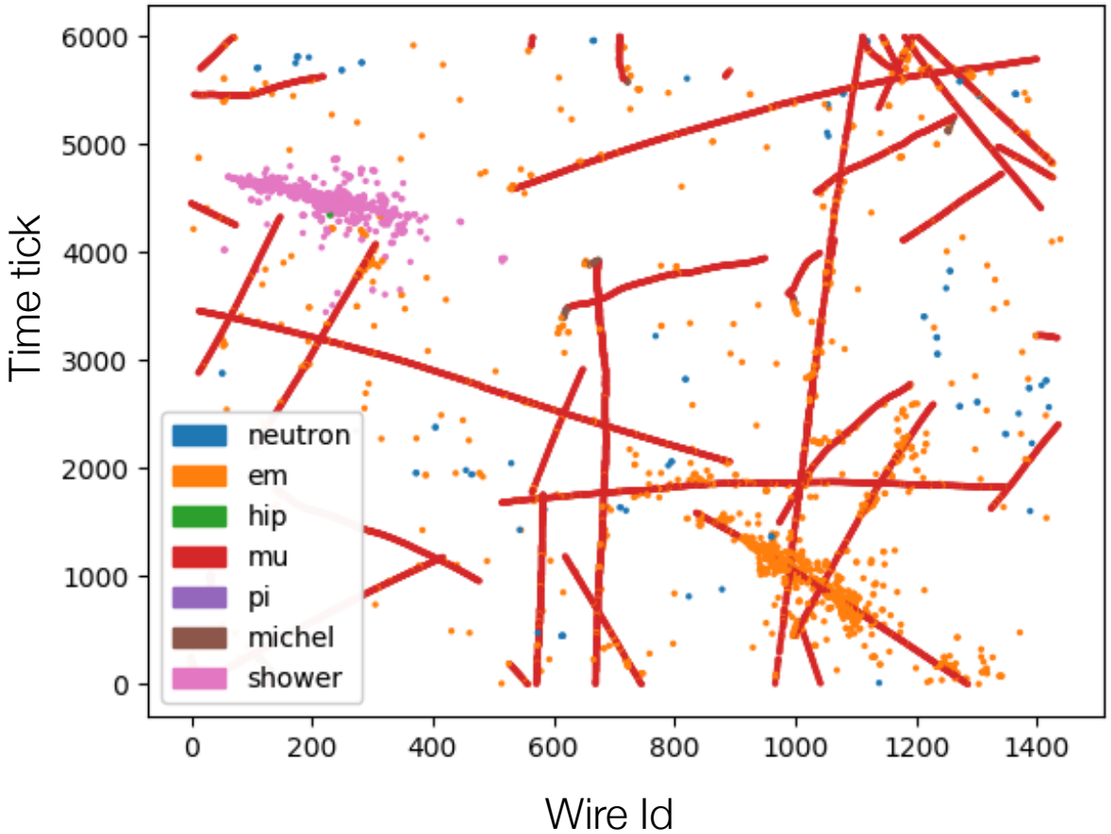
Carlos Sarasty Segura
1st April 2020
DRA meeting

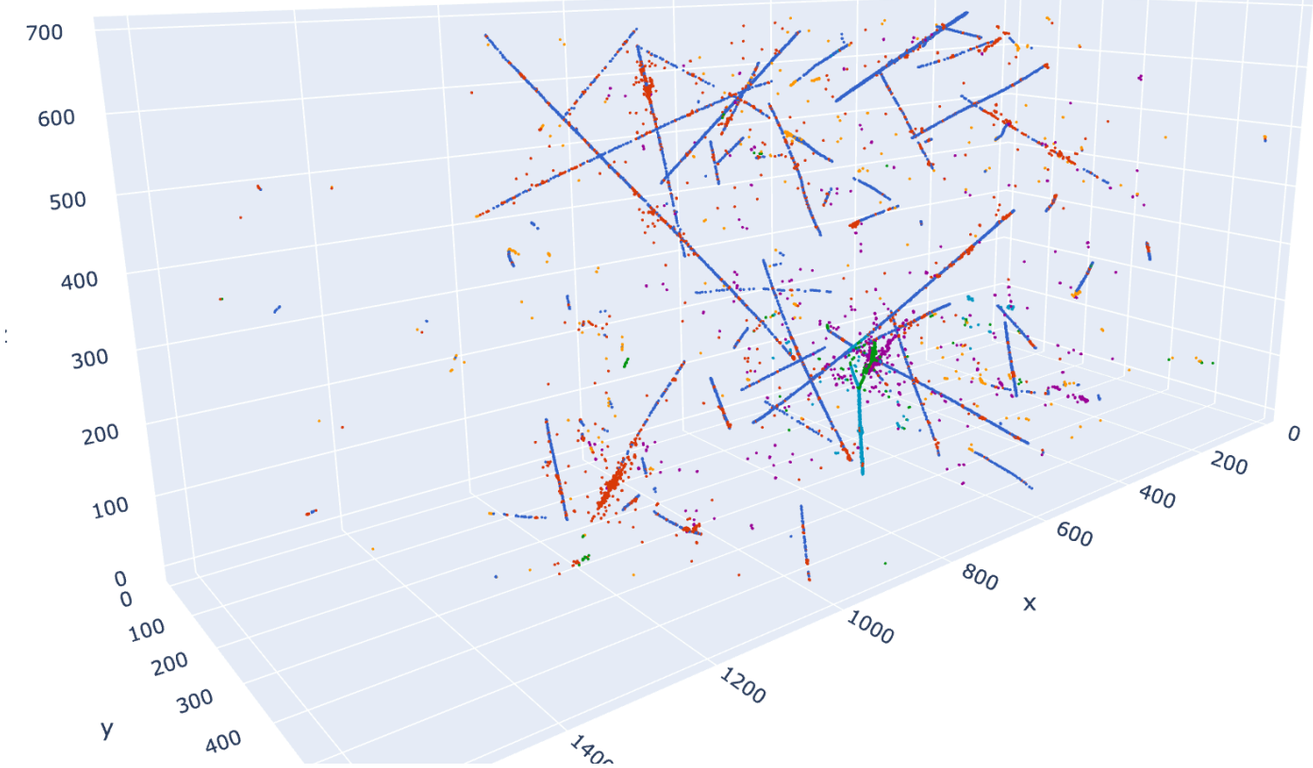
Outline

- Definition of the ground truth
- Training using 2D samples
- Training using 3D samples
- Summary

Semantic Segmentation

- Goal: Apply sparse CNNs for the task of semantic segmentation at a pixel level in ProtoDUNE



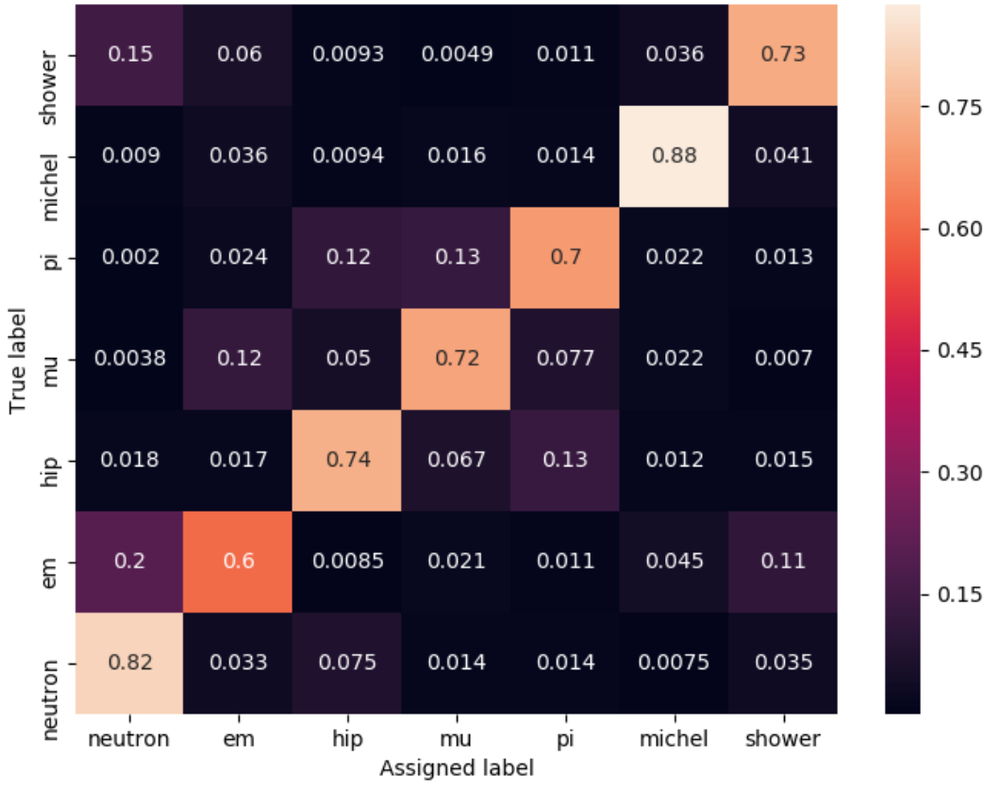


Ground Truth - First Version

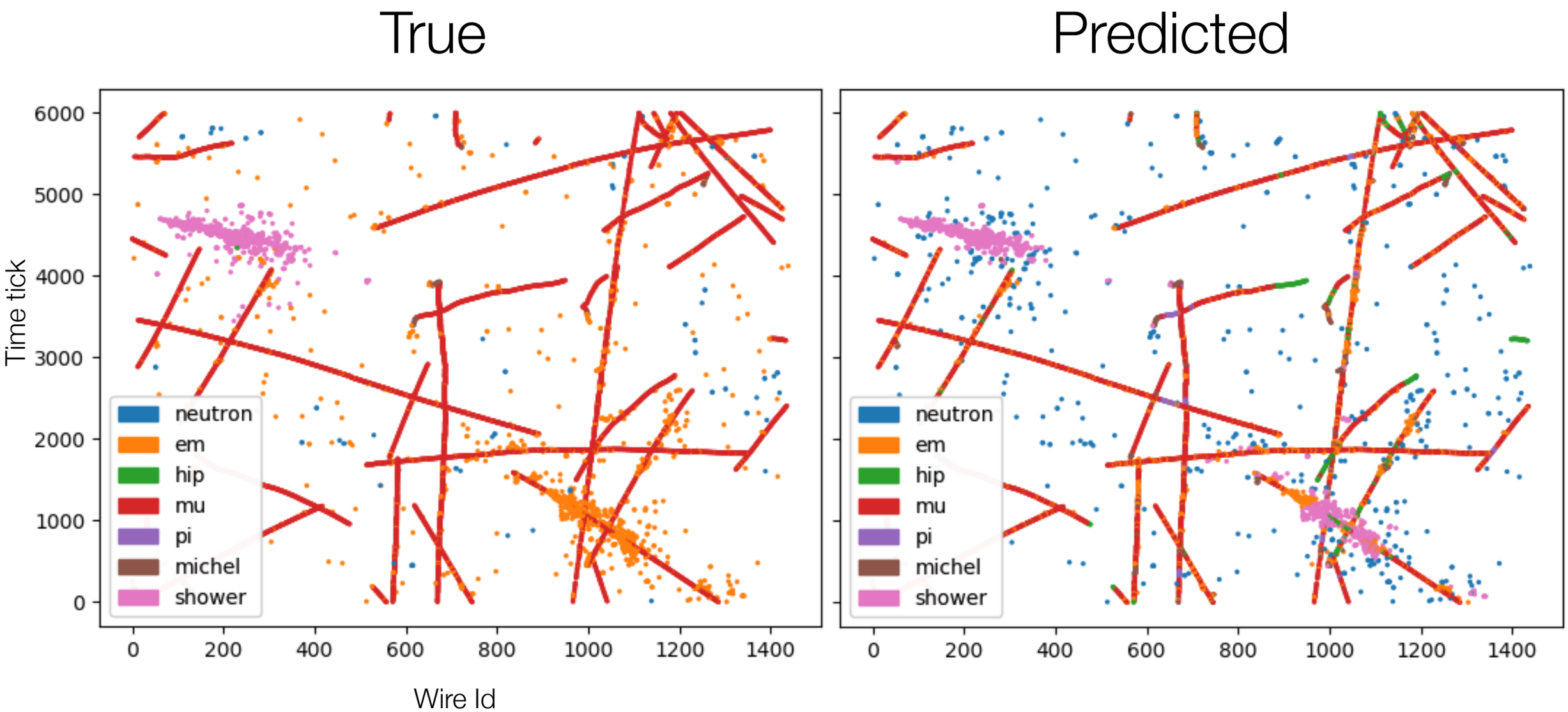
- Classify each pixel into 7 different classes for supervised learning
  - **MIP** → Two classes → **muons & pions**
  - **HIP** → **Protons, kaons & nuclei**
  - **Showers** → Induced by electromagnetic particles such as e^- and e^+
  - **Michel electrons** → From the decay of muons
  - **Electromagnetic activity** → Electrons from hard scattering, and low energy e^-
  - **Neutrons**
  - Record the fraction of energy deposited by each class per pixel
  - https://indico.fnal.gov/event/20144/session/17/contribution/93/material/slides/0.pdf

Supervised Learning

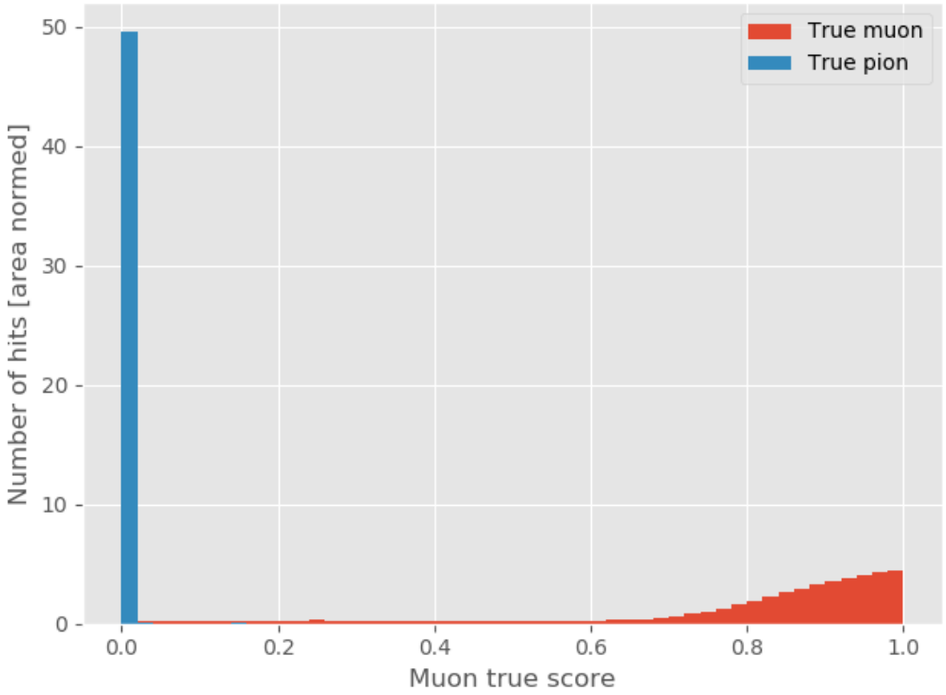
- The dataset consist of about 100.000 2D image samples of up to 6000 px split into 95% and 5% for train and test respectively
- 1 feature → Integrated cargue

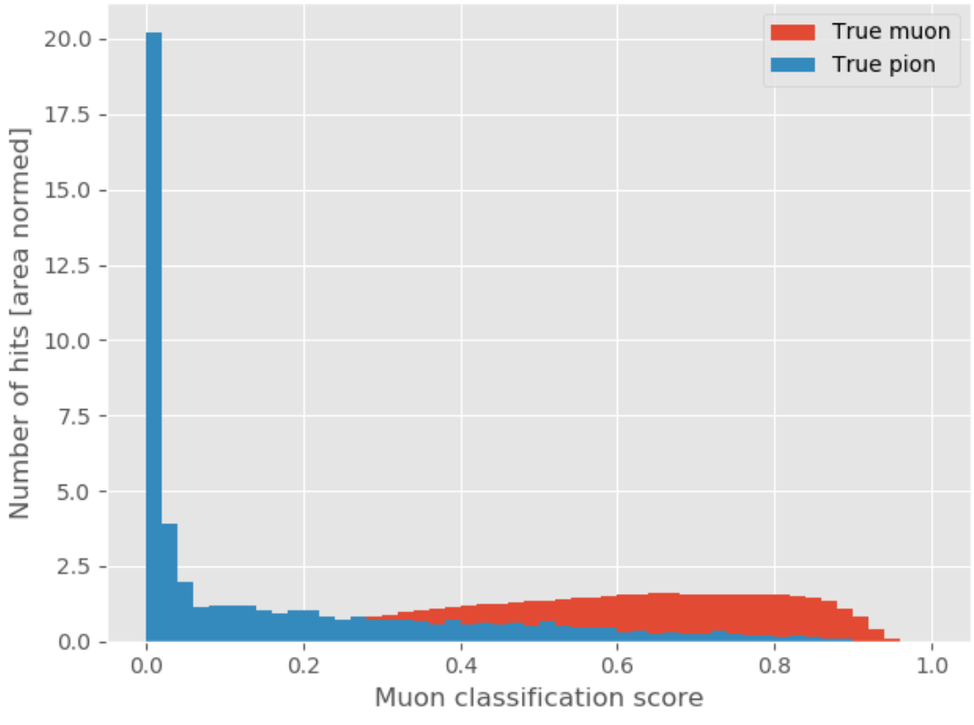


Event Display Example

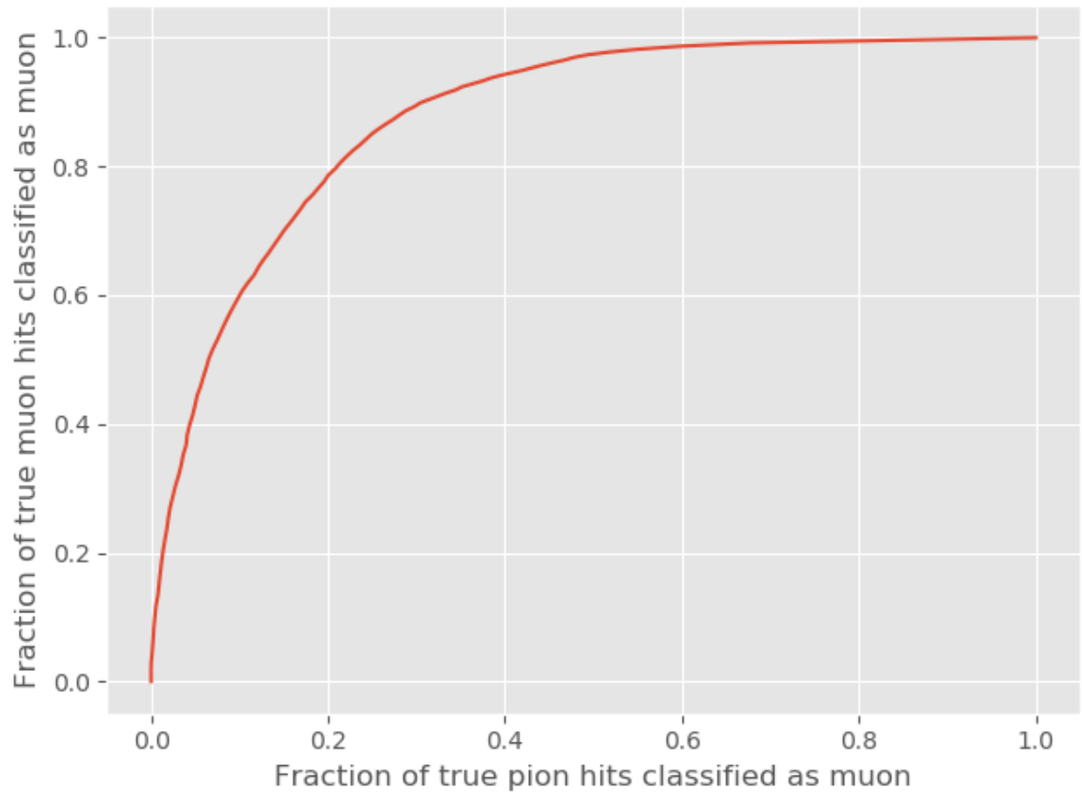


Muon-Pion Separation





ROC curve

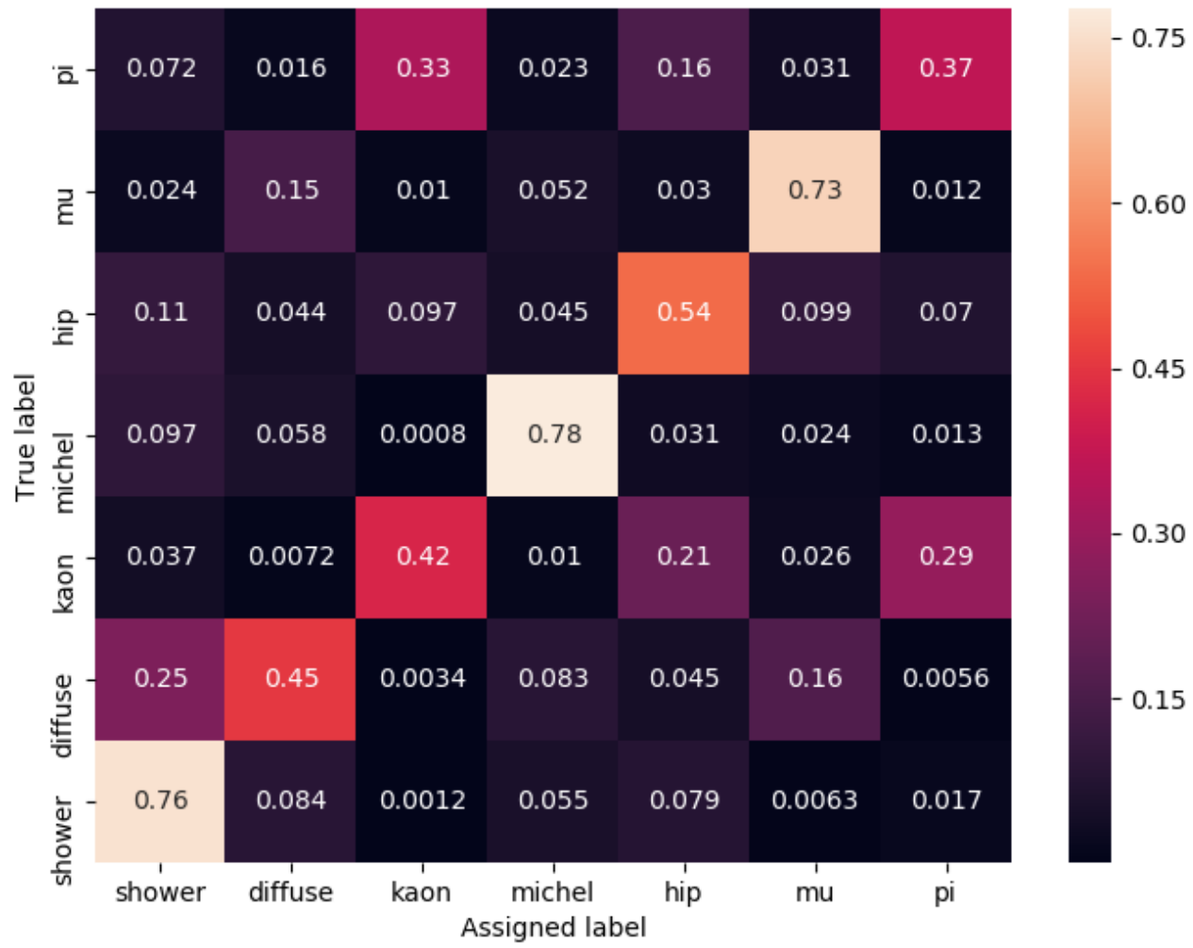


Moving to 3D

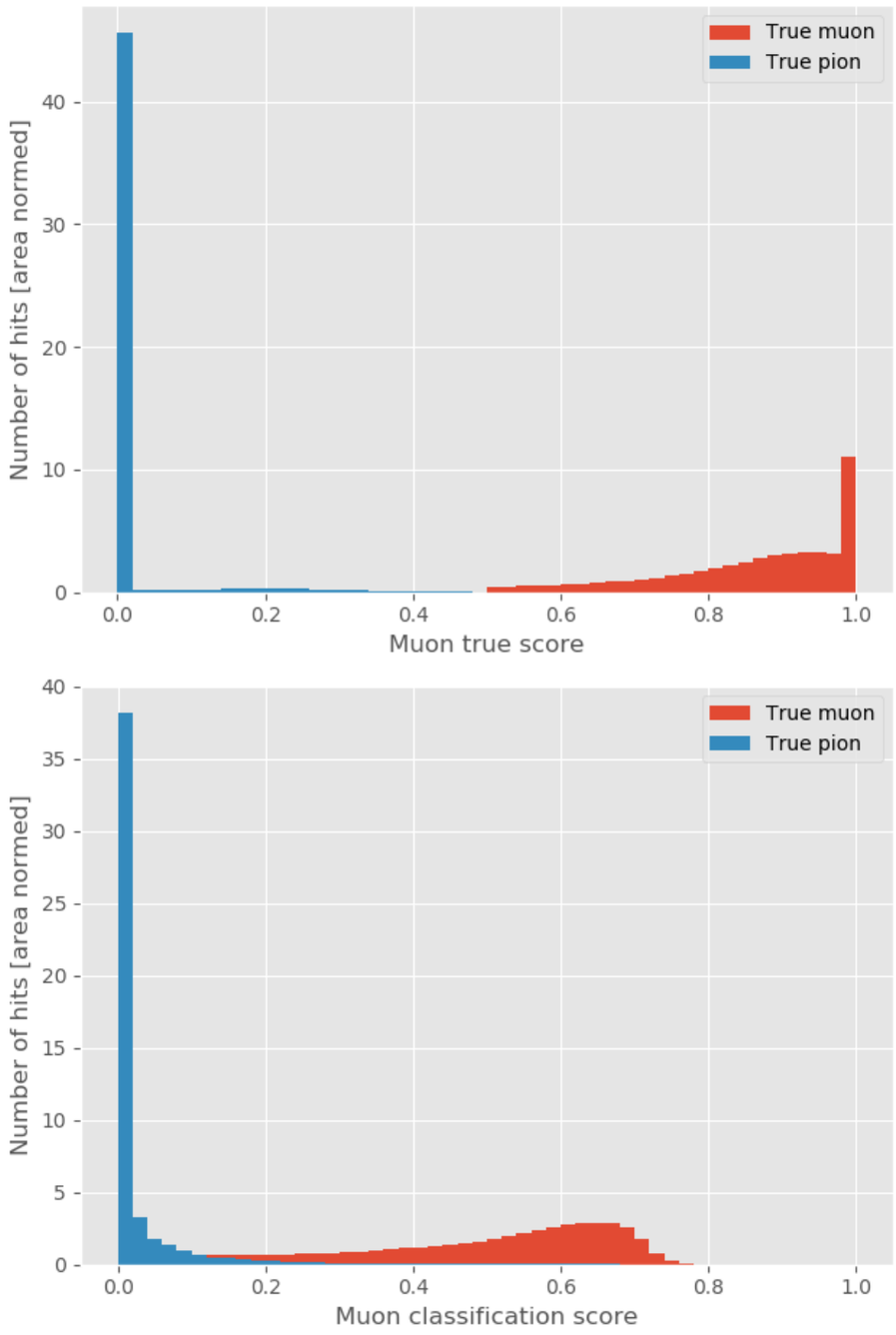
- **Ground truth:**
  - Modify the ground truth definition to separate **kaons** from the **hip** class.
  - Merge **neutrons & EM activity** into 1 class
- **Features:**
  - Increase the number of features from **1** to **7** (3 coordinates per hit, integrated charge per plane per voxel, number of hits per voxel)
- **Issues:**
  - Low statistics for the kaon class → only **5%** of files contain kaons

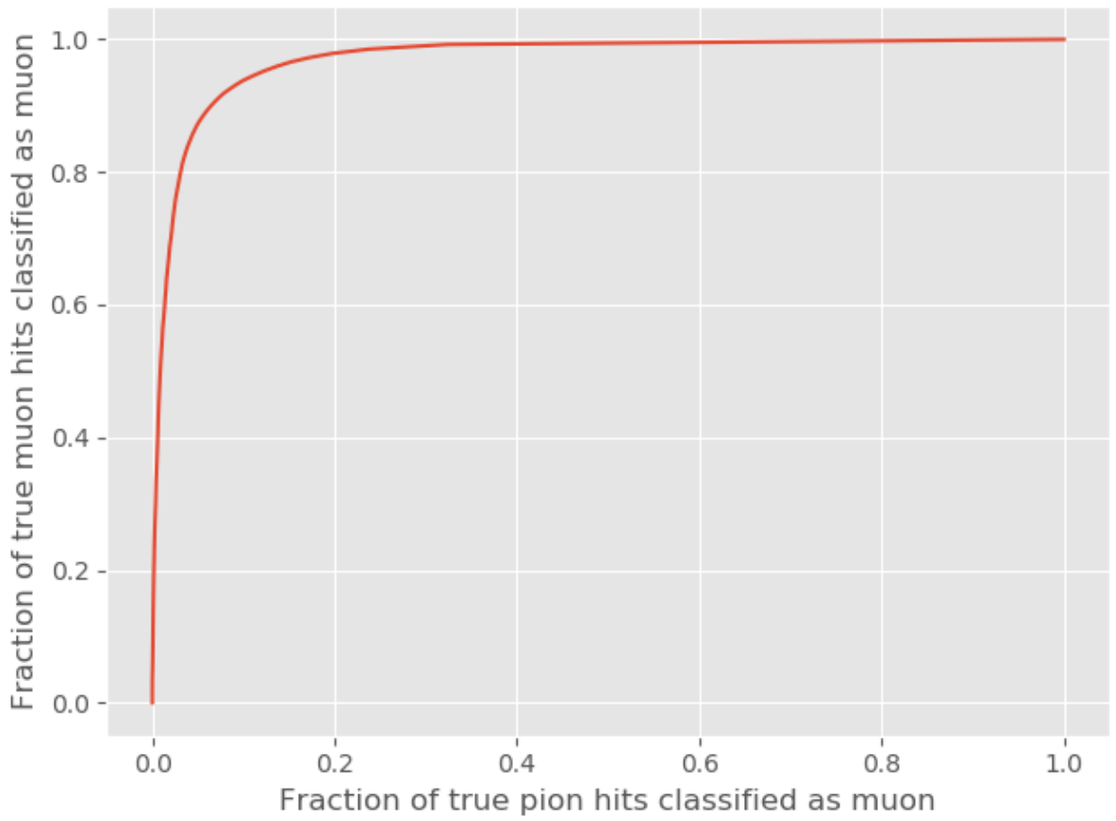
Supervised Learning

- The dataset with kaons consist of **3943** 3D images split into 95% and 5% for train and test



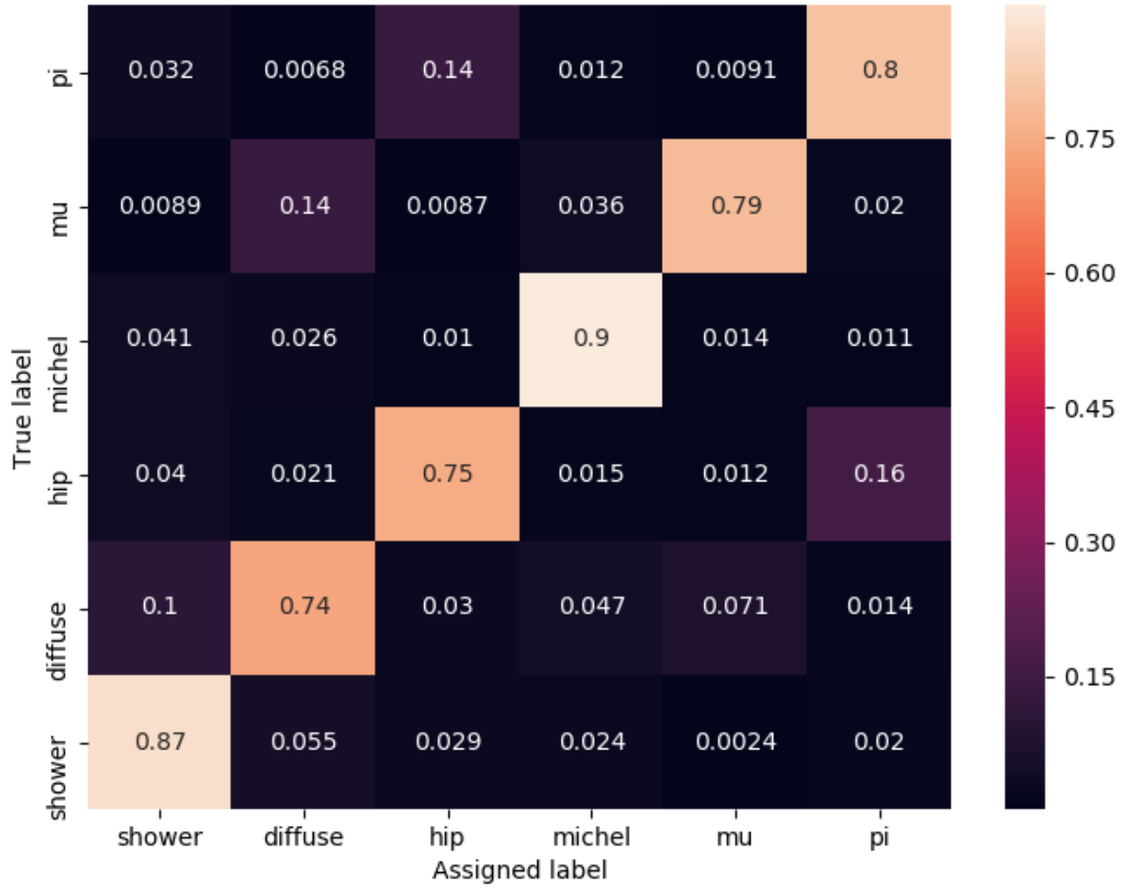
Muon-Pion Separation



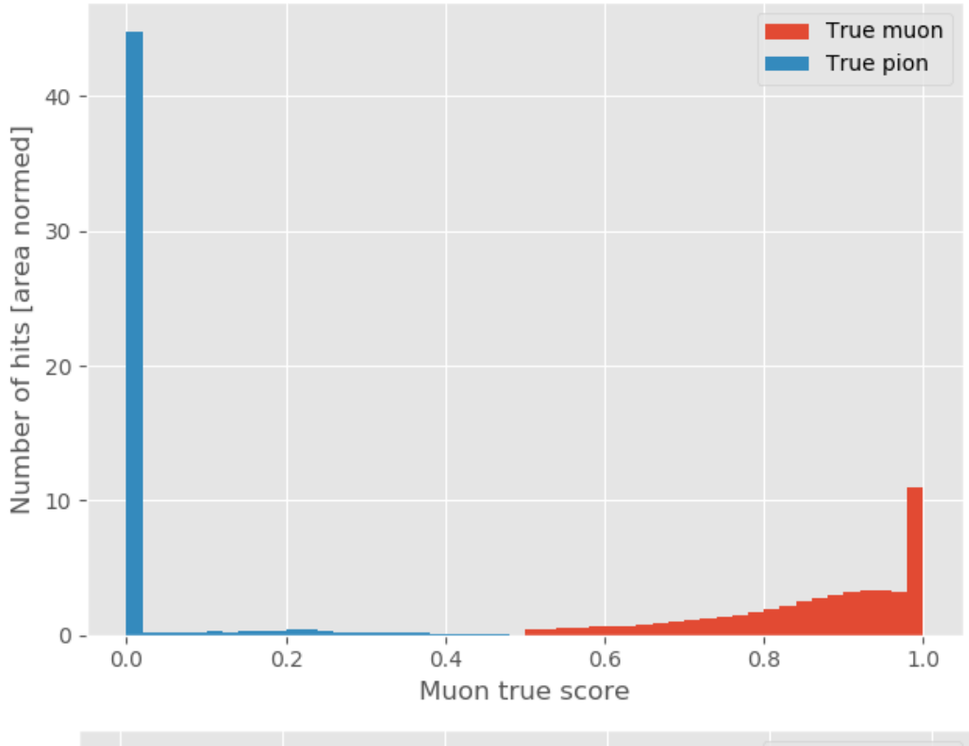


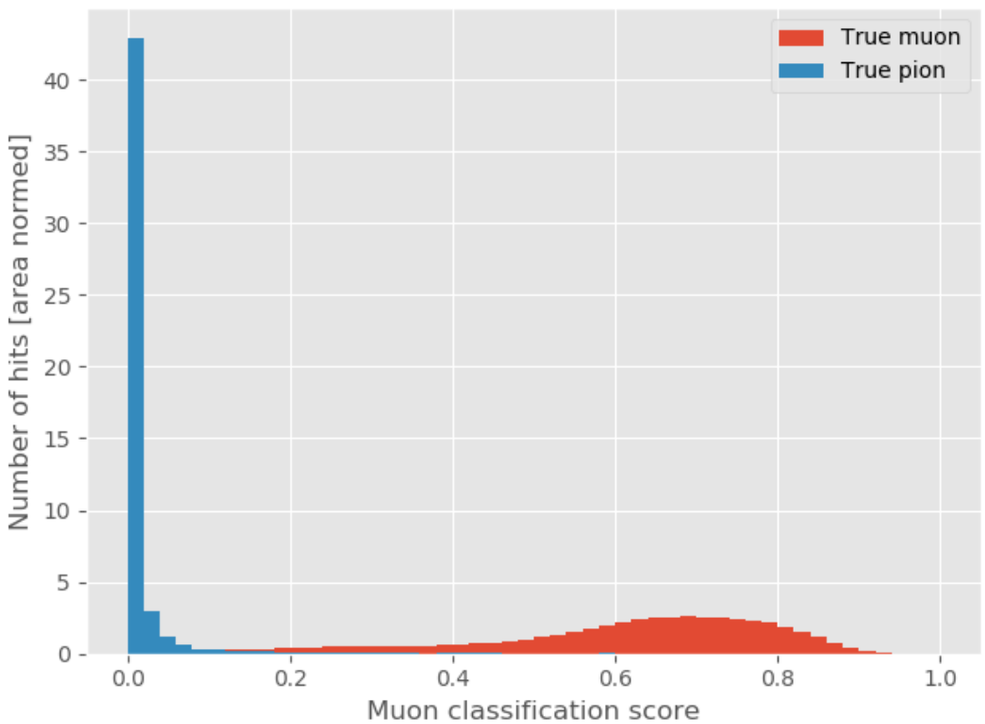
Second case

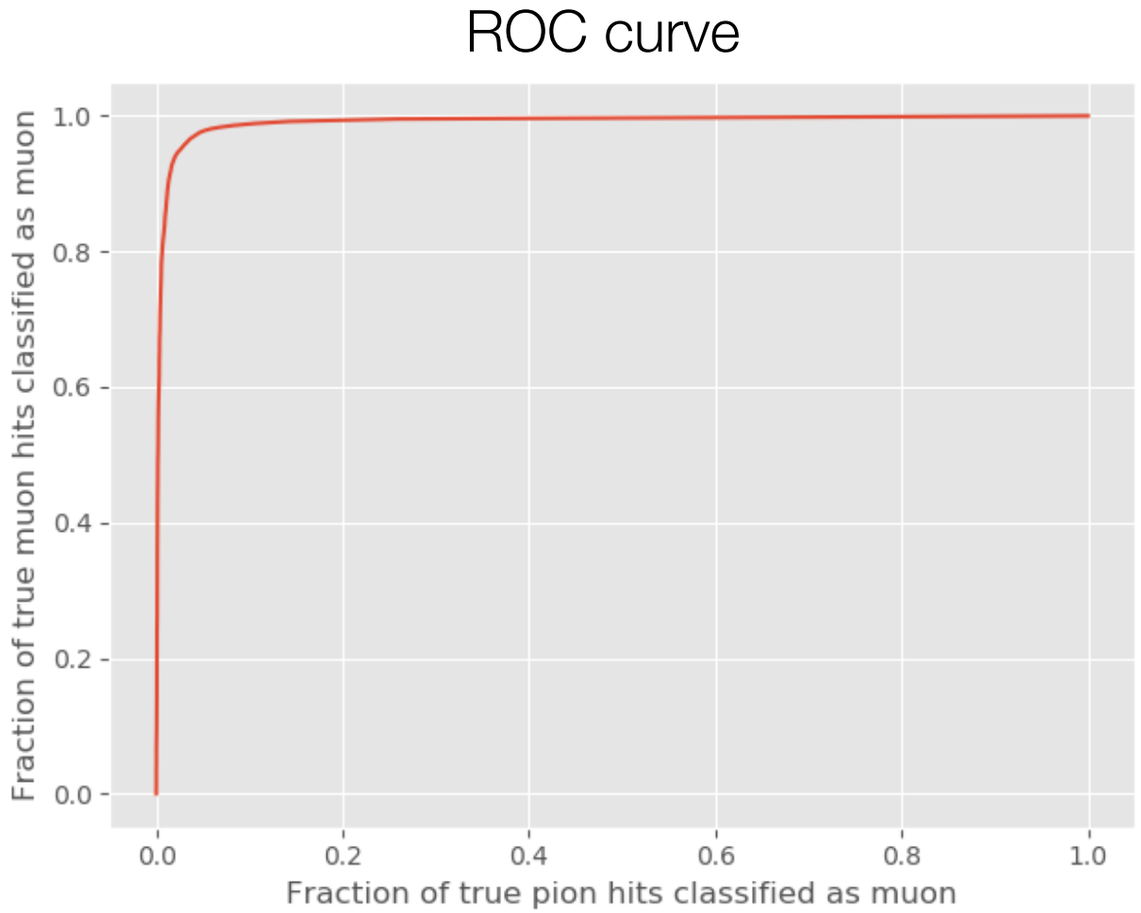
- **Ground truth:**
  - Merge **kaons** back in with **protons** into **hip** class
- **Dataset:** Consist of **70k** 3D images



Muon-Pion Separation

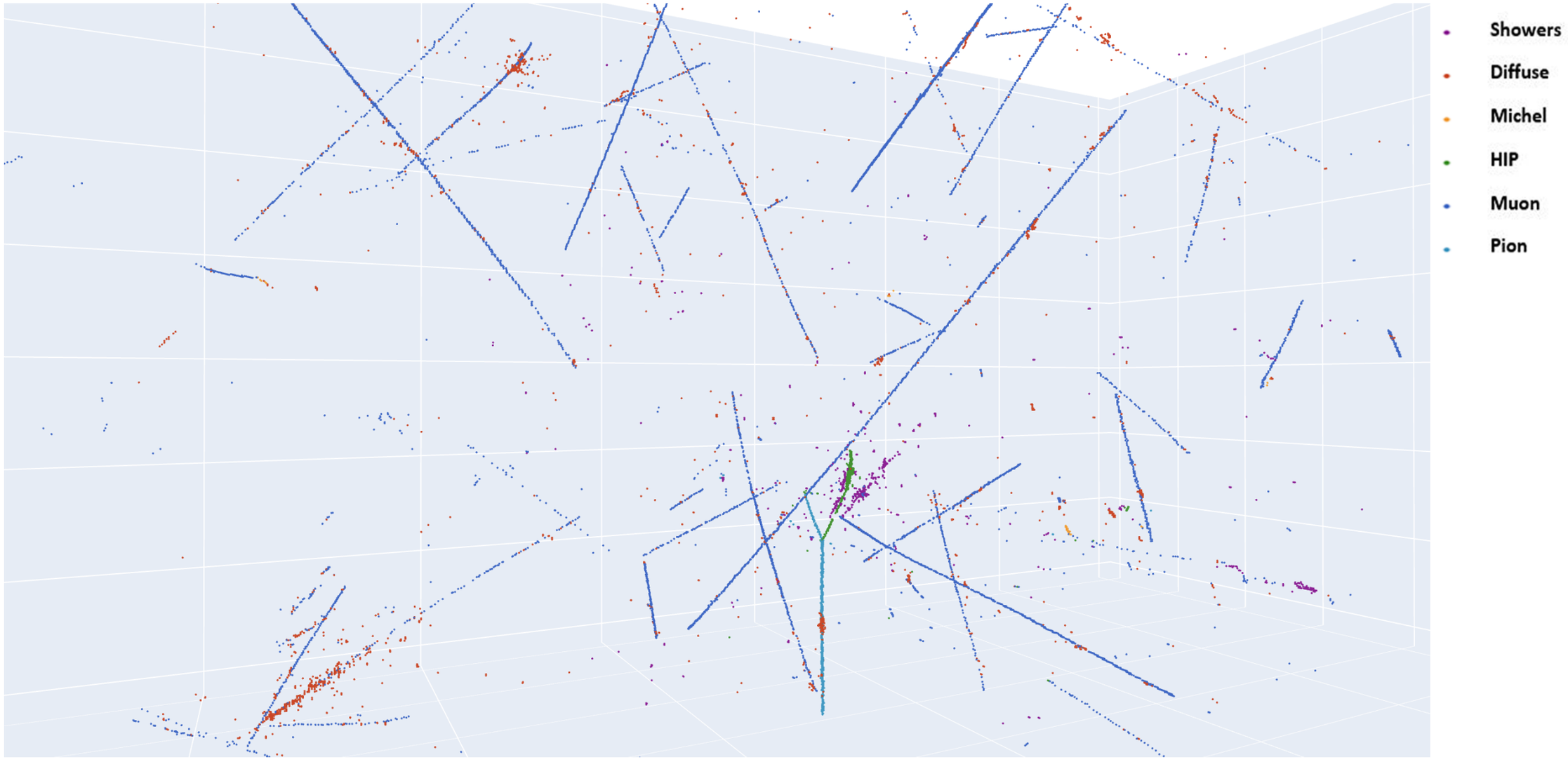




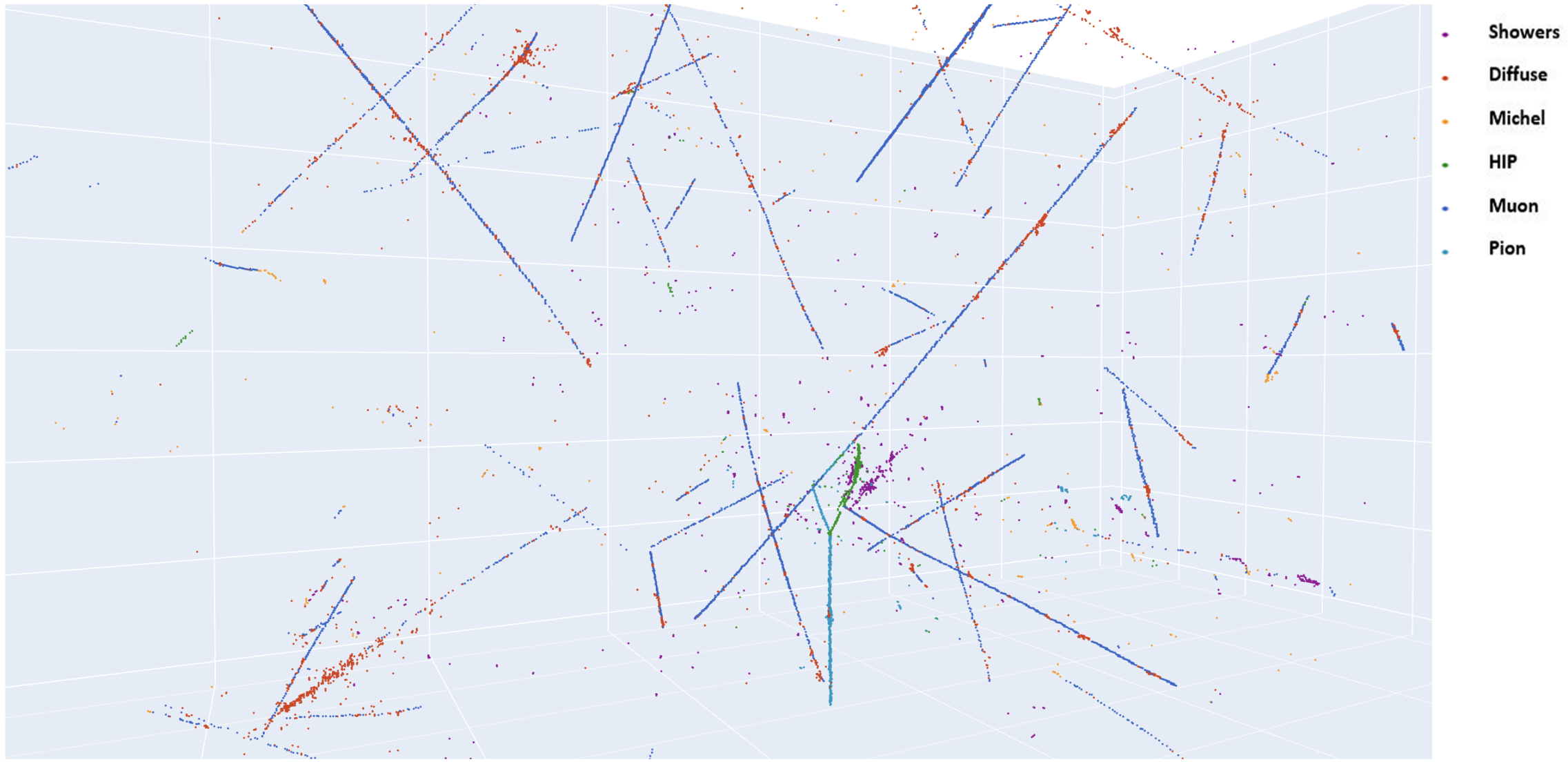


University of CINCINNATI

Event display - True



Event display - Predicted



TO DO:

- Modify the ground truth:
  - Include Delta rays as a separate class.
  - Separate electron and photon showers
- Retrain and test the model for electron and photon separation.

Summary

- We have trained the network using different definitions of the ground truth and different datasets
- The performance of the network using 3D samples is significantly better than the 2D case
- A training using kaons as a separate class can be possible with a bigger dataset
- Comments and suggestions are more than welcome
- Thanks! :)

Backup slides

Ground Truth

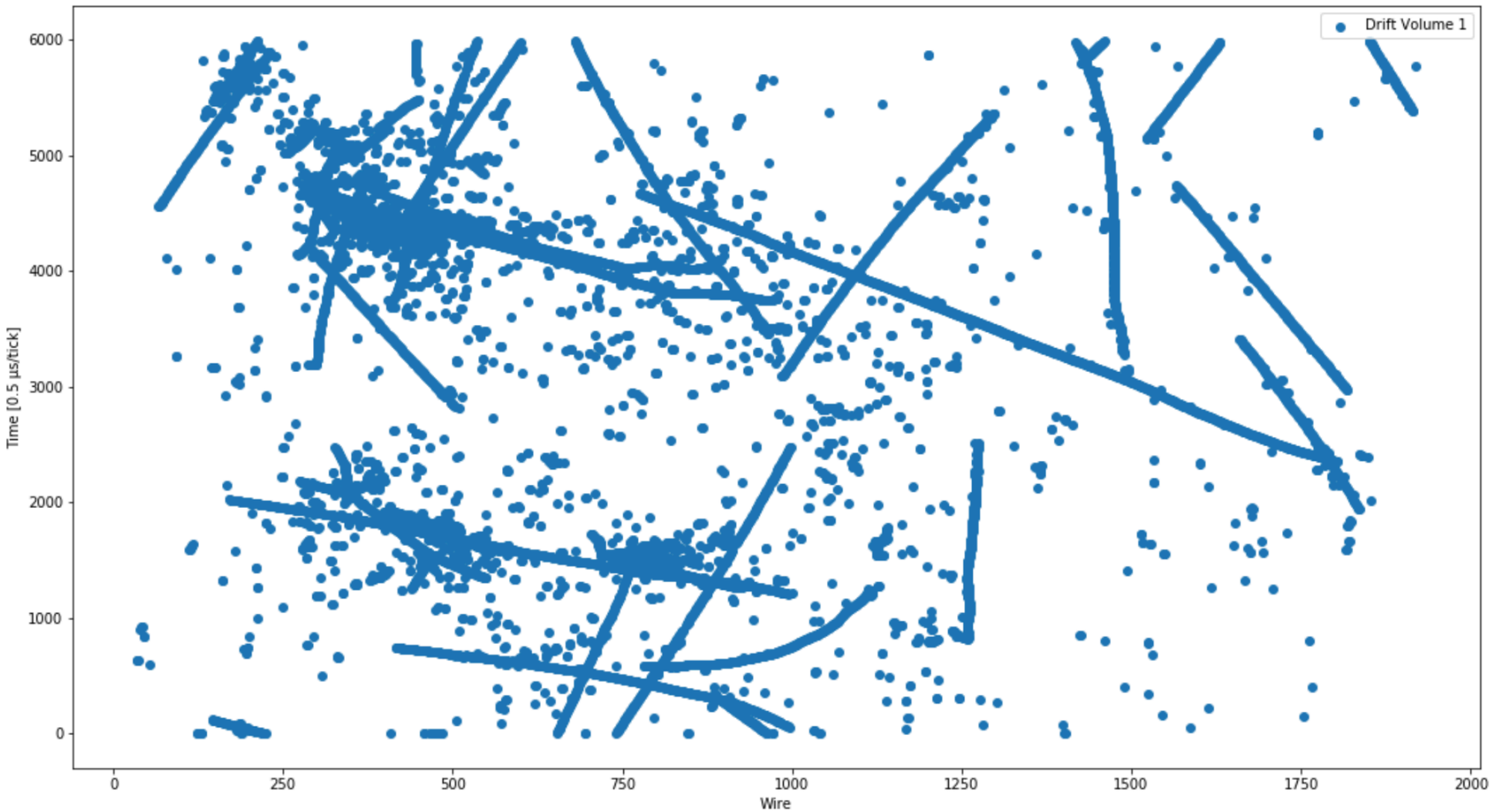
- The first approach to distinguish the different classes of particles is based on the **pdg** and **track Id** information
- **Geant4** also provides valuable information of the physical process of a simulated particle and its parent. This information is useful to characterize **Michel electrons**
  - Non-primary electron
  - Electron's parent is a muon
  - Same with positrons.
- **Neutrons**
  - **Check the process** → n-capture , neutron Inelastic scattering

Ground Truth

- **EM showers** and **EM activity**
- In the MC Truth the information of secondary and tertiary particles from showers is thrown away → ***shower daughters*** are tagged with the negative track ID of the parent particle
- Identify all particles that belong to the same track ID
- Set a threshold in the number of hits → nhits > 10 ~ 5cm
- Any other $e^{+/-}$ will be labeled as **EM activity**

Ground Truth

Drift Volume 1



University of CINCINNATI

Ground Truth

Event display - True